SELECTION SYSTEM FOR EVOLVING PROTEASES AND PROTEASE-CLEAVAGE SITES

FIELD OF THE INVENTION

[0001] The present disclosure pertains to the field of molecular biology. More specifically, the present disclosure provides a fusion protein, which can be used to develop orthogonal proteases/protease cleavage-sites, methods using said fusion protein, as well as variant protease cleavage-sites of bdSUMO and variant proteases of bdSENP1.

BACKGROUND OF THE INVENTION

[0002] Recombinant protein expression and purification are fundamental for modern biochemistry and structural biology as well as for the production of proteins and protein complexes for practical applications.

[0003] The most exploited protein expression system is the bacterium *E. coli* because of the ease of its genetic manipulation, low cost of biomass production and fast growth kinetics. *E. coli* fails, however, to introduce typical eukaryotic post-translational modifications and in many cases also to fold eukaryotic proteins properly. Expression in eukaryotic hosts is then an alternative.

[0004] Affinity tags facilitate the purification of recombinant proteins considerably. Poly histidine tagged proteins, for example, can be purified in a standardized manner, namely by binding to a Ni(II) chelate matrix, washing off contaminants and desorption by imidazole.

[0005] The maltose-binding protein (MBP), thioredoxin (TRX), or NusA exemplify another category of tags, namely those that enhance the level of expression, solubility and successful folding of the fused target protein, in particular if *E. coli* is used as an expression host. Tags may interfere with the target protein's function or the intended application. It is therefore often desired to remove them from the target protein. Tag removal is typically implemented by (I) modular fusions of an N-terminal tag (or tags), an intervening protease-cleavage site (PCS) and the protein of choice; (II) by treatment of the purified fusion protein with an appropriate PCS-specific protease and (III) separation of the target protein from the cleaved tag and the protease.

[0006] The "affinity capture and proteolytic release strategy" is a particularly efficient implementation of affinity chromatography. A tag-PCS-target protein fusion is bound through the tag to an affinity matrix. However, elution is not achieved by disengaging the tag from the matrix, but by cleavage of the PCS module. This procedure combines the specificities of the affinity matrix with that of the protease and therefore yields far purer protein preparations than affinity chromatography alone. Furthermore, it is time-saving and simplifies the workflow by eliminating the need for separate tag cleavage and tag removal.

[0007] An extension of this strategy even allows to select for the presence of several subunits within a given protein complex. This requires that two or more subunits are equipped with distinct (and non-cross reacting) affinity tags and with distinct PCSs, and that two or more rounds of affinity capture and proteolytic release are performed in succession (Frey & Görlich 2014b). Furthermore, it requires that the used proteases are orthogonal in their specificities, i.e. that each protease cleaves only "its" PCS and leaves the others intact.

[0008] One type of tag-cleaving proteases recognizes short linear peptides as PCSs, examples being: thrombin, Factor Xa, enterokinase, human rhinovirus 3C protease, or TEV protease. These suffer, however, from one or several of the following problems: poor specificity leading also to degradation of the target protein, poor substrate turnover, poor activity at low temperature (4° C.), remaining undesired residues on the target protein, or difficult production of the protease (reviewed in Yan et al. 2009).

[0009] Proteases that recognize ubiquitin-like modifiers (Ubls), such as SUMO, NEDD8 or Atg4 as a PCSs, overcome these difficulties (Malakhov et al. 2004; Frey & Görlich 2015; Frey & Görlich 2014a). They recognize the 3D fold of the cognate Ubls and cleave behind a Gly-Gly motif (Phe-Gly or Tyr-Gly in the case of Atg8), accept all amino acids (except for Pro) at the P_1' position, and they show a tremendous substrate turnover even at temperatures as low as 0° C. (Frey & Görlich 2014a). Previously used Ubl-proteases include: *S. cerevisiae* Ulp1 (also called SENP1, cleaving SUMO-fusions), *Brachypodium distachyon* bdSENP1 (also cleaving SUMO-fusions), *Brachypodium distachyon* or *Salmon salar* NEDP1 (cleaving NEDD8-fusions), *Xenopus laevis* Atg4B (cleaving Atg8-fusions).

[0010] An additional advantage of the use of Ubls as PCSs is that they have (like MBP, NusA or thioredoxin) a strong expression- and fold-enhancing effect. They can, however, not be used in eukaryotic expressing systems, because the endogenous Ubl-proteases cause a premature tag-cleavage. So far, there was only one attempt to solve this problem, namely the SUMOstar system (Patent US 2012/0065106 A1). SUMOstar is derived from yeast SUMO (scSUMO) and carries the R64T and R17E mutations that disrupt the Ulp1-binding interface, while the SUMOstar protease is a D451S, T452G and E455S-mutated version of the wild type yeast Ulp1 protease. SUMOstar protease cleaves wild type scSUMO as well as the SUMOstar fusions with comparable efficiency. It is thus not orthogonal to the parental protease.

[0011] The purpose of this invention was to create a generic in vivo selection system to evolve protease/PCS pairs to enhanced and novel specificities, and in particular to orthogonality to a reference pair. A further aim was to evolve SUMO mutants that can be used as a stable fusion tag in any eukaryotic systems and that is resistant to cleavage SUMO-specific proteases from yeast, plants, human, amphibians, and insects as well as to cleavage by the SUMOstar protease. A final aim was to evolve a variant bdSENP1 protease, which (I) efficiently cleaves the new bdSUMO mutant(s), but not wild type scSUMO, human SUMO (i.e. hsSUMO2) or SUMOstar-fusions and (II) causes no toxicity when expressed in a prokaryotic or eukaryotic host.

SUMMARY OF THE INVENTION

[0012] The present disclosure relates to a selection system for evolving proteases and protease-cleavage modules to novel specificities. We describe a SUMO protein mutant ($bdSUMO^{Mut1}$) as a novel protease-cleavage module, which is highly resistant to cleavage by previously described SUMO proteases and thus allows the stable expression of $bdSUMO^{Mut1}$-fusion proteins in eukaryotic hosts. We further describe a SUMO protease mutant ($bdSENP1^{MutB}$), which leaves SUMOstar and wild type SUMO fusion proteins intact but cleaves $bdSUMO^{Mut1}$-fusion proteins efficiently and thus enables tag-removal as well as protein